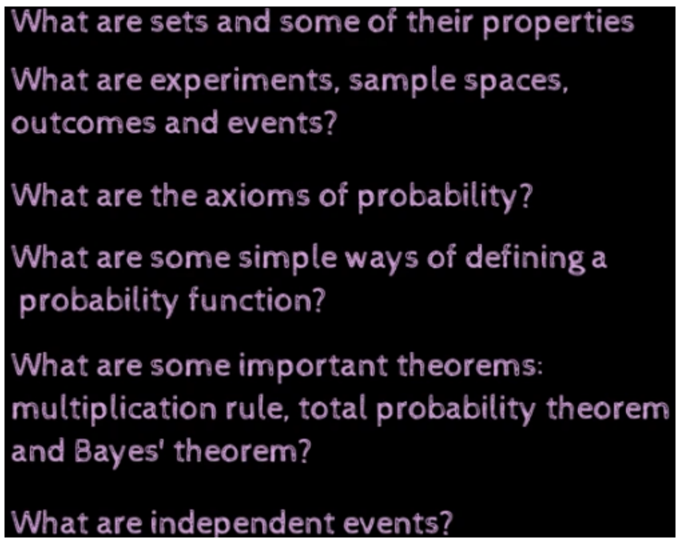
# Week 15 : Sample Spaces and Events

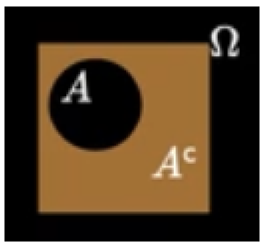
| | |
|---|---|
| pending tasks | |
| type | |

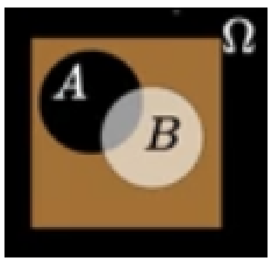
**Learning Objectives:**




## An Overview of Set Theory

- A set is a collection of elements. The order does not matter. A compact notation can be used to represent large sets. e.g. S = {a, e, i, o, u} here, $a \in S$ but $b \notin S$.
- All the elements of a subset can be found in the parent set. e.g. set of vowels S is a subset of the set of alphabet A or, $S \subset A$.
- Two sets are equal if they are subsets of each other, this can be stated as $A = B\ iff\ A \subset B\ and\ B \subset A$.
- Every set of interest is a subset of the universal set e.g. set of all aces (A) is a subset of the deck of cards $\Omega$, or $A \subset \Omega$.
- A complement of a set A includes all elements of the universal set $\Omega$ that do not belong the set. $A^c = \{x : x \in \Omega\ and\ x \notin A\}$

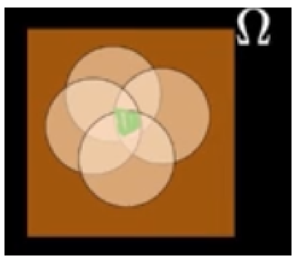



black,white,gray for union; only gray for intersection.

- Union of sets A and B includes all the elements belonging to the two sets. $A \cup B = \{x : x \in A\ or\ x \in B\}$
- Intersection of two sets A,B contains all elements that belong to both A and B. $A \cap B = \{x : x \in A\ and\ x \in B\}$.

- Union of n set can be given by $x \in A_1 \cup A_2 \cup A_3 ... \cup A_n\ iff\ x \in A_i\ for\ some\ i.$
- Intersection of n set is given by $x \in A_1 \cap A_2 \cap A_3 ... \cap A_n\ iff\ x \in A_i \forall i$




All white for union; green for intersection.

## Properties of Set Operations

- **Commutativity**

  $A \cup B = B \cup A$
  $A \cap B = B \cap A$

- **Associativity**

  $A \cup (B \cup C) = (A \cup B) \cup C$
  $A \cap (B \cap C) = (A \cap B) \cap C$

- **Distributive Laws**

  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

  **Proof ( first one)**

  $x \in A \cap (B \cup C)$
  $\implies x \in A \; and \; x \in (B \cup C)$
  $\implies x \in A \; and \; B \; or \; x \in A \; and \; C$

- **DeMorgan's Laws**

  $(A \cup B)^c = A^c \cap B^c$
  $(A \cap B)^c = A^c \cup B^c$

  **Proof (first one)**

  $x \in (A \cap B)^c$
  $\implies x \notin A \cap B$
  $\implies x \notin A \; or \; x \notin B$
  $\implies x \in A^c \cup B^c$

- **Countable v/s uncountable Infinite Sets**

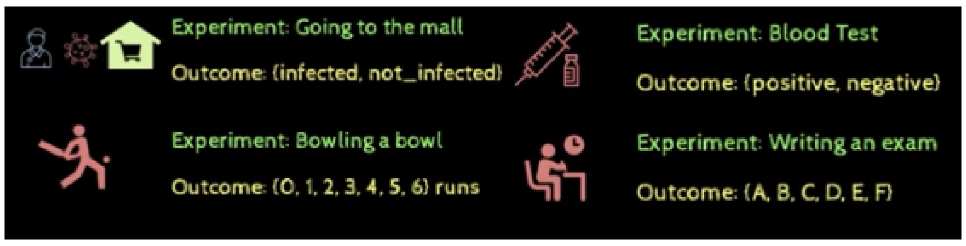
  Set of all real numbers has infinite elements and is uncountable, whereas the set of integers has infinite elements and is countable. An infinite set is countable if there is a 1-1 correspondence between the elements of the set and the set of positive integers.

## Experiments & Sample Spaces

*def experiment - (t = 1:55)*

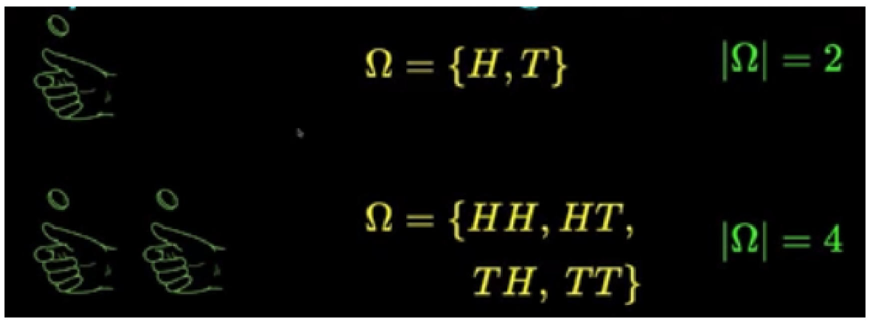
*def sample space - (t = 2:40)*

- Procedure that can be repeated infinite timed and has a well defined set of outcomes is called an **experiment** or a **trial.**



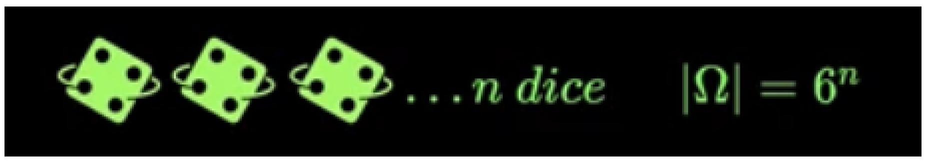

Examples of experiments.

- **Sample space** is the set of all possible outcomes of an experiment. Its elements are **mutually exclusive** and **collectively exhaustive.**
- The outcome of every trial is uncertain but the set of outcomes is certain.
- For example, Tossing a coin is a classic example of an experiment. The sample space $\Omega$ differs with number of coins being tossed simultaneously.




Example of tossing coins.

- Rolling fair dice is another such example where there are 6 possible outcomes on a single roll.




Example of rolling n dice.

- The higher the number the outcomes (equally likely) of an experiment, lesser the probability of any single outcome. The experiments can also have continuous outcomes.

## Events of an Experiment

*def event - (t = 0:30)*

- The set of outcomes of an experiment is called an **event.** It is a subset of the sample space. For instance, in the sample space of double coin toss (with four possible outcomes) the event that first toss coressponds to a head (which is a subset of the sample space) is {HH, HT}
- Union of two events A and B would include all the events that occur in A and B $(A \cup B)$ . Likewise, the intersection would include elements that are common to both A and B $(A \cap B)$. Compliment of an event A $(A^c)$ would include all the elemnts of the sample space not present in A, thus, $A \cup A^c = \Omega$ and $A \cap A^c = \emptyset$.
- Union of multiple events (e.g. A,B and C) is such that the outcome occurs in atleast on of A, B or C. The intersection would imlpy that the outcome be a part of all three sets.
- Two events A and B are said to be disjoint if they cannot occur simultaneously, i.e., $A \cap B = \emptyset$. $A$ and $A^c$ would be one such example. In general, events $A_1, A_2, A_3.., A_n$ are said to be mutually disjoint or pairwise disjoint if $A_i \cap A_j = \emptyset \forall i, j$ such that $i \neq j$.
- If union of mutually disjoint sets is equal to the universal set then they are said to partition the sample space (e.g. if $A \cup B \cup C = \Omega$ , then A,B,C partition the sample space)

## Axioms of Probability

- Probability of an event A , P(A) is called the probability function. The properties of probability functions if given by axioms of probability.

  **Axiom 1 (non-negativity) :** $P(A) \geq 0 \forall A$. No event can have a negative chance of occurrence.

  **Axiom 2 (normalization) :** $P(\Omega) = 1$. The probability of the sample space must be 1.

  **Axiom 3 (finite additivity) :** If $A_1, A_2..., A_n$ are mutually disjoint, then $P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$. This implies that the probability of larger events can be computed from smaller events. Fo6 instance, each of the six possible outcomes of a die roll is known, then the probabilities other events (such as chance of an odd number) can be easily calculated.

## Some Properties of Probability

**Property 1:** $P(A) = 1 - P(A^c)$

This follows from axiom 2 and $A \cup A^c = \Omega$.

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$$
$$\implies P(A) = 1 - P(A^c)$$

**Property 2:** $P(A) \leq 1$

From axiom1, in $P(A) = 1 - P(A^c)$ , $P(A^c)$ must be greater than or equal to 0. This implies that the value of P(A) must be greater than 1, proving the property 2.

**Property 3:**$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Writing B as $(B \cap A^c)$ in $(A \cup B)$ensures that the sets A and B are disjoint. This would enable the application of axiom 3.

$P(A \cup B) = P(A \cup (B \cap A^c)$
$= P(A) + P(B \cap A^c)$

P(B) can be written as $P((B \cap A^c) \cup (B \cap A))$ and since both are disjoint sets, the additivity property can be applied.

Thus, $P(B) = P(B \cap A^c) + P(B \cap A)$

Therefore, $P(B \cap A^c) = P(B) - P(B \cap A)$

using this, $P(A \cup B) = P(A) + P(B) - P(B \cap A)$

**Property 4: The sum of probabilities of all outcomes is equal to 1**

Since outcomes form partiti ons of sample space $(\Omega = A_1 \cup A_2 .. \cup A_n)$ and all the sets are disjoint ( from axiom 3, $P(\Omega) = P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$). Here RHS is equal to 1, this implies that LHS is also equal to 1 i.e. $\sum_{i=1}^{n} P(A_i) = 1.$

**Property 5:** $P(\emptyset = 0)$

The probability of the universal set is 1. and $P(\Omega) = P(\Omega \cup \emptyset)$, which are disjoint.

Thus, $1 = P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$

since, $P(\Omega) = 1$, $P(\emptyset) = 0.$

## Designing Probability Functions (as relative frequency)

- Probability of an event $A_i$ can be thought of as the fraction of times the event occurs when the experiment is repeated a **large number** of times.

$$P(A_i) = \frac{no.of\ times\ outcome\ is\ in\ A_i}{no.of\ times\ the\ experiment\ was\ repeated}$$

- $P(A_i)$ is a ratio of 2 positive numbers. This implies $P(A_i) \geq 0$, thus, satisfying the first axiom of probability. Since the outcome always has to be in the sample space $P(\Omega) = 1$, which satisfies the second axiom.
- Let $A_1\ and\ A_2$ be two disjoint sets with number of events $k_1\ and\ k_2$ respectively and the total number of events (from both the sets) be n. This implies that that we need to check if $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ :

$$P(A_1 \cup A_2) = \frac{k_1 + k_2}{n}$$
$$= \frac{k_1}{n} + \frac{k_2}{n} = P(A_1) + P(A_2)$$

Therefore, the third axiom is also satisfied.

- The axioms of probability are about all possible events (outcomes being a part of it).  Any event can be represented as a union of some disjoint outcomes. Thus, if the frequencies of outcomes are known then the probability of any event can be computed.
- The sample from which the probabilities were estimated should be drawn from the same population on which we are interested in making inferences.

## Designing Probability Functions (equally likely outcomes)

- Deriving a formula for computing the probability of events of an experiment with n equally likely outcomes:

$E$ : any event with k outcomes (disjoint)

Thus, $E$ can be expressed as a union of 'k' singleton events(outcomes):

$P(E) = \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}$

'k' is the number of outcomes in event X divided by the total number of outcomes in $\Omega$.

i.e. the probability function $P(X) = \frac{\text{no. outcomes in X}}{\text{no. outcomes in } \Omega}$

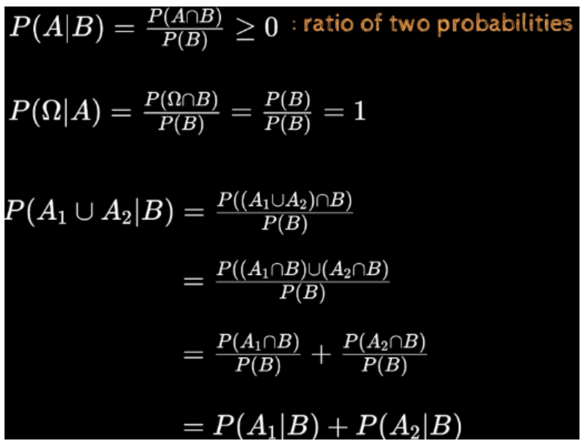
- Here, the first axiom of probability is satisfied ( $P(A_i) \geq 0$)since is the ratio of two numbers. The second one is also satisfied as $P(\Omega) = 1$ in the above equation. It also satisfies the third axiom of probability.

## Conditional Probabilities

- $P(A|B)$ is called the conditional probability of event A given the event B. It is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

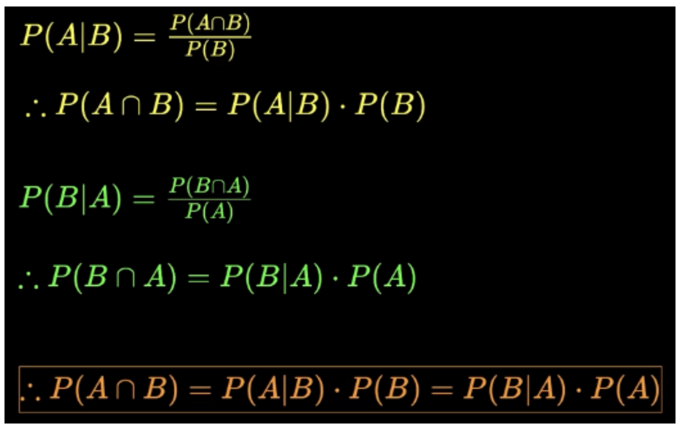
- Conditional probabilities also satisfy the axioms of probability.




Axioms of probability.

## The Multiplication Principle

- It is also called as chain rule of probability.



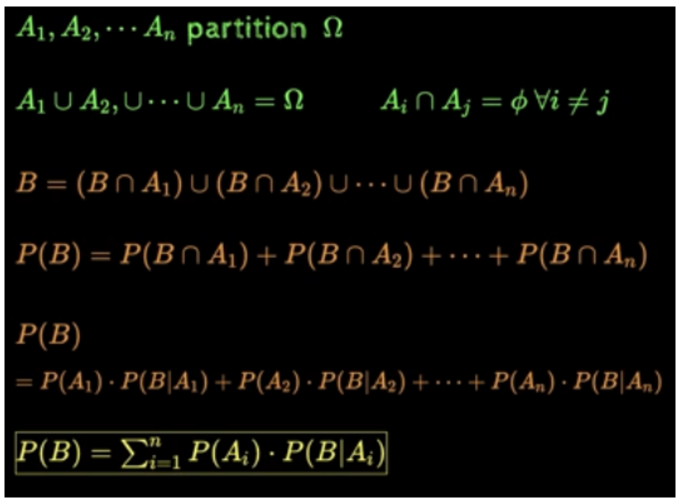

Chain rule of probability.

- For n events, $P(A_1 \cap A_2 \cap .. \cap A_n) = P(A_1) \prod_{i=2}^{n} P(A_i|A_1 \cap ..A_{i-1})$.

## Total Probability Theorem

A set B (partition of $\Omega$) which intersects 'n' disjoint sets (the 'n' sets together form $\Omega$) can be expressed as union of its parts. The probability of the set B can be found using the total probability theorem as

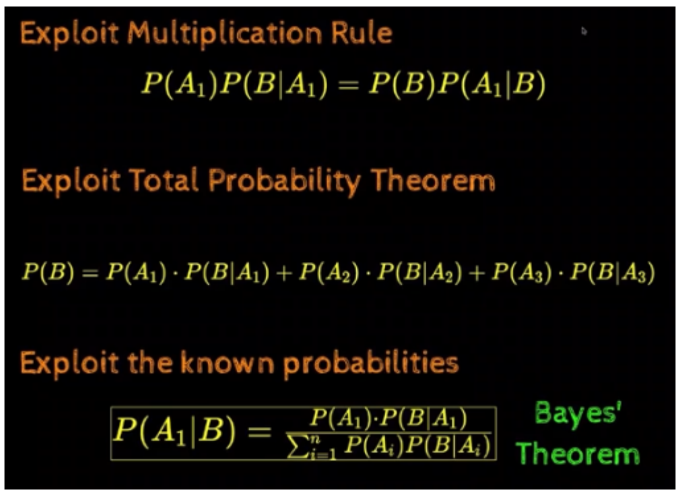
$$P(B) = \sum_{i=1}^{n} P(A_i).P(B|A_i)$$




Total Probability theorem.

## Bayes' Theorem

Bayes' Theorem is about updating existing belief in a event when the result of another event is known.




## Independent Events

*def independent events - (t=5:24)*

*def pairwise independence - (t=12:00)*

*def mutual independence - (t=12:29)*

- Event probabilities are updated using Bayes' only when they are dependent. This does not apply to independent events.
- Two events A and B are independent if $P(A \cap B) = P(A).P(B)$
- Events $A_1, A_2..A_n$ are said to be pairwise independent if for any two events $P(A_i \cap A_j) = P(A_i).P(A_j) \forall i \neq j$.
- Mutual independence is more strict than pairwise independence in the sense that $A_1, A_2..A_n$ are said to be (mutually) independent if for all subsets $I \subset (1, 2, 3, ...n)$ of n events, the condition, $P(\cap_{i \in I} A_i) = \prod_{i=1}^{n} P(A_i)$ holds.

## Summary

- There is an element of chance (randomness) associated with the everyday activities, therefore the certainty of an event cannot be precisely determined.
- Sets can be infinite countable/uncountable.
- The probability of an event can be computed as the sum of the probabilities of the disjoint outcomes contained in the event.
- Any event can be represented as a union of some disjoint outcomes. Thus, if the frequencies of outcomes are known then the probability of any event can be computed.
- Conditional probability P(A|B) is given by, $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- By the chain rule of probability, $P(A \cap B) = P(A|B).P(B) = P(B|A).P(A)$.
- The total probability theorem states that an event B that is a formed from elements of disjoint sets (that make up $\Omega$) is given by $P(B) = \sum_{i=1}^{n} P(A_i).P(B|A_i)$.

- Using the Bayes' theorem we can update the known probability as and when more evidence for related events are obtained. This can be achieved using $P(A_1|B) = \frac{P(A_1).P(B|A_1)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$. This isn't used if the events are independent.